

33. The method of claim 1, further comprising denaturing the recombinant nucleic acid, and contacting the recombinant nucleic acid with at least one additional nucleic acid from the oligonucleotide set.

34. The method of claim 1, further comprising denaturing the recombinant nucleic acid, and contacting the recombinant nucleic acid with at least one additional nucleic acid produced by cleavage of a parental nucleic acid encoded by the at least one parental character string.

35. The method of claim 1, further comprising denaturing the recombinant nucleic acid, and contacting the recombinant nucleic acid with at least one additional nucleic acid produced by cleavage of a parental nucleic acid encoded by the at least one parental character string, which parental nucleic acid is cleaved by one or more of: chemical cleavage, cleavage with a DNase and cleavage with a restriction endonuclease.

36. The method of claim 1, wherein the parental character string encodes one or more nucleic acid corresponding to one or more of protein or gene selected from: EPO, insulin, a peptide hormone, a cytokine, epidermal growth factor, fibroblast growth factor, hepatocyte growth factor, insulin-like growth factor, an interferon, an interleukins, a keratinocyte growth factor, a leukemia inhibitory factor, oncostatin M, PD-ECSF, PDGF, pleiotropin, SCF, c-kit ligand, VEGEF, G-CSF, an oncogene, a tumor suppressor, a steroid hormone receptor, a plant hormone, a disease resistance gene, an herbicide resistance gene, a bacterial gene, a monooxygenases, a protease, a nuclease, and a lipase.

37. The method of claim 1, wherein the set of oligonucleotides comprises one or more oligonucleotide member between about 20 and about 60 nucleotides in length.

38. The method of claim 1, further comprising selecting the recombinant nucleic acid for a desired trait or property, thereby providing a selected recombinant nucleic acid.

39. The method of claim 38, further comprising recombining the selected recombinant nucleic acid with one or more of: a homologous nucleic acid, and an oligonucleotide member from the set of oligonucleotides.

40. The method of claim 1, further comprising selecting the recombinant nucleic acid for a desired trait or property, thereby providing a selected recombinant nucleic acid, wherein the desired trait or property is selected in an in vivo selection assay or a parallel solid phase assay.

5 **41.** The method of claim 1, further comprising selecting the recombinant nucleic acid for a desired trait or property, thereby providing a selected recombinant nucleic acid, wherein the desired trait or property is selected in an in vitro selection assay.

42. The method of claim 1, further comprising deconvolution of the recombinant nucleic acid.

10 **43.** The method of claim 1, further comprising sequencing or cloning the recombinant nucleic acid.

44. The method of claim 1, wherein the recombinant nucleic acid is synthesized in vitro by assembly PCR.

15 **45.** The method of claim 1, wherein the recombinant nucleic acid is synthesized in vitro by error-prone assembly PCR.

46. The method of claim 1, wherein the parental character strings, or oligonucleotide sets are selected in a computer.

47. A method of making character strings, the method comprising:

- 20 a) providing a parental character string encoding a polynucleotide or polypeptide;
 b) providing a set of oligonucleotide character strings of a pre-selected length that encode a plurality of single-stranded oligonucleotide sequences comprising sequence fragments of the parental character string, and complement thereof;
 c) creating a set of derivatives of the parental sequence comprising sequence variant strings, the set comprising a plurality of mutations, having one mutation per variant string.

25 **48.** The method of claim 47, wherein a plurality of the plurality of single-stranded oligonucleotide sequences are overlapping in sequence.

49. The method of claim 47, further comprising applying one or more genetic operator to the parental character string, or to one or more of the oligonucleotide character

strings, wherein the genetic operator is selected from: a mutation of the parental character string, or one or more of the oligonucleotide character strings, a multiplication of the parental character string, or one or more of the oligonucleotide character strings, a fragmentation of the parental character string, or one or more of the oligonucleotide character strings, a crossover
5 between any of the parental character string or one or more of the oligonucleotide character strings, or an additional character string, a ligation of the of the parental character string, or one or more of the oligonucleotide character strings, an elitism calculation, a calculation of sequence homology or sequence similarity of an alignment comprising the parental character string, or one or more of the oligonucleotide character strings, a recursive use of one or more
10 genetic operator for evolution of character strings, application of a randomness operator to the parental character string, or to one or more of the oligonucleotide character strings, a deletion mutation of the parental character string, or one or more of the oligonucleotide character strings, an insertion mutation into the parental character string, or one or more of the oligonucleotide character strings, subtraction of the of the parental character string, or one or
15 more of the oligonucleotide character strings with an inactive sequence, selection of the of the parental character string, or one or more of the oligonucleotide character strings with an active sequence, and death of the parental character string, or one or more of the oligonucleotide character strings.

50. The method of claim 47, further comprising:

- 20 d) providing a set of overlapping character strings of a pre-defined length that encode both strands of the parental character string sequence; and,
e) synthesizing sets of single-stranded oligonucleotides according to the step (c) and (d).

51. The method of claim 50, further comprising:

- 25 f) assembling a library of recombinant nucleic acids by assembly PCR from the single-stranded oligonucleotides.

52. A library made by the method of claim 51.

53. The method of claim 51, further comprising:

- g) selecting or screening the library for one or more recombinant polynucleotide having
30 a desired property.

54. The method of claim 52, further comprising:
h) deconvoluting the sequence of the one or more selected polynucleotide.

55. The method of claim 50, wherein the sequence of the one or more selected polynucleotide is deconvoluted by sequencing the selected polynucleotide, or by digesting the one or more selected polynucleotide.

56. The method of claim 50, wherein the sequence is deconvoluted by positional deconvolution of the one or more selected polynucleotide

57. The method of claim 50, further comprising reiterative shuffling or selection of the library of recombinant nucleic acids.

58. A method of facilitating recombination between two or more divergent nucleic acids, the method comprising:
aligning parental character strings corresponding to the divergent nucleic acids, thereby identifying regions of sequence identity and regions of sequence diversity;
defining a diplomat character string which is intermediate in sequence between the parental character strings;
synthesizing at least a portion of the diplomat sequence to produce a diplomat nucleic acid; and,
recombining a mixture of selected nucleic acids comprising the parental nucleic acids, or fragments thereof, and the diplomat nucleic acid.

59. The method of claim 58, wherein the diplomat nucleic acid is synthesized by synthesizing a plurality of overlapping oligonucleotides corresponding in sequence to the diplomat sequence, hybridizing the overlapping oligonucleotides, and incubating the overlapping oligonucleotides with a polymerase.

60. The method of claim 58, further comprising synthesizing a pool of oligonucleotides corresponding to one or more of the parental character strings, which pool of oligonucleotides is present in the mixture of selected nucleic acids.

61. The mixture of selected nucleic acids produced by the method of claim 60.

62. A method of generating and recombining nucleic acids, the method comprising:

inputting a plurality of amino acid sequence character strings into a digital system;

reverse translating the amino acid character strings in the digital system into a plurality
5 of nucleic acid character strings, wherein reverse translated nucleic acid sequences are selected
for one or more of: species codon bias in a selected expression host, and optimized sequence
similarity between the plurality of nucleic acid character strings; and,

synthesizing one or more sets of oligonucleotides corresponding to one or more reverse
translated nucleic acid sequences.

10 **63.** The method of claim 62, further comprising hybridizing members of the
one or more oligonucleotide sets to each other, or to a set of fragmented nucleic acids which
encodes one or more amino acid polymer corresponding to one or more of the amino acid
sequence character strings.

64. The method of claim 63, further comprising elongating one or more
15 resulting hybridized nucleic acids with a polymerase.

65. The method of claim 62, further comprising fragmenting one or more
resulting elongated nucleic acids and hybridizing the resulting secondary fragmented nucleic
acids with each other or with members of the one or more oligonucleotide sets, or with a set of
primary fragmented nucleic acids which encodes one or more amino acid polymer
20 corresponding to one or more of the amino acid sequence character strings.

66. A method of optimizing activity of a nucleic acid, the method comprising:
parameterizing a set of nucleic acids or proteins to provide a set of multidimensional
datapoints;

extrapolating one or more postulated multidimensional datapoint from the set of
25 multidimensional datapoints; and,

converting the postulated multidimensional datapoint to a new character string
corresponding to a postulated nucleic acid nucleic acid or protein.

67. The method of claim 66, comprising synthesizing the postulated nucleic
acid or protein.

68. The method of claim 66, further comprising principle component analysis of the set of multidimensional datapoints.

69. The method of claim 66, comprising shuffling the postulated nucleic acid, or a subsequence thereof, with an additional nucleic acid.

5 **70.** The method of claim 66, wherein the set of nucleic acids or proteins is parameterized by correlating each residue of the nucleic acid or protein to a matrix of numeric indicators.

10 **71.** The method of claim 70, wherein the matrix is graphically represented as a tetrahedron, having an assigned origin at the center of the tetrahedron, with each corner represented as a numeric representation, with each residue of a nucleic acid being positioned at a different corner, thereby producing the matrix of numeric indicators.

72. The method of claim 66, comprising correlating each multidimensional datapoint with an output vector to identify a relationship between a matrix of dependent Y variables and a matrix of predictor X variables.

15 **73.** The method of claim 72, wherein the correlation is performed by partial least square projections to latent structures analysis.

20 **74.** The method of claim 66, wherein each multidimensional datapoint comprises more than one different parameter, wherein the parameters are plotted against each other in n dimensional hyperspace, said n dimensional hyperspace comprising at least one dimension for each parameter.

75. A method of providing a library of recombinant nucleic acids which is enriched for a sequence of interest and selecting the library, the method comprising:

producing an initial library of at least about 10^6 recombinant nucleic acids, which initial library of recombinant nucleic acids comprises at least about 10^5 different member types,

25 which 10^5 different member types are non-identical;

hybridizing the library to one or more population of nucleic acids, which one or more population of nucleic acids correspond to one or more subsequences in the different library members;

isolating members of the library which hybridize to the one or more populations of nucleic acids, thereby enriching the library of nucleic acid for members which hybridized to the one or more population of nucleic acids; and,

selecting members of the resulting enriched library for one or more property of interest.

5 **76.** The method of claim 75, wherein the initial library has between about 10^9 and 10^{12} members.

77. The method of claim 75, wherein the one or more population of nucleic acids is fixed to a solid substrate.

10 **78.** The method of claim 77, wherein the solid substrate comprises one or more of: a column matrix material and a nucleic acid chip.

79. The method of claim 75, wherein the initial library is produced by recombining one or more homologous nucleic acids.

80. The enriched library produced by the method of claim 75.

15 **81.** The method of claim 75, wherein the initial library is produced by:
providing a plurality of parental character strings corresponding to a plurality of nucleic acids, which character strings, when aligned for maximum identity, comprise at least one region of similarity and at least one region of heterology;

aligning the character strings;

20 defining a set of character string subsequences, which set of subsequences comprises
subsequences of at least two of the plurality of parental character strings;

providing a set of oligonucleotides corresponding to the set of character string subsequences;

annealing the set of oligonucleotides; and,

25 elongating one or more member of the set of oligonucleotides with a polymerase,
thereby producing the initial library of nucleic acids.

82. A method of generating a library of biological polymers, the method comprising:

generating a diverse population of character strings in a computer, which character strings are generated by alteration of pre-existing character strings; and,

synthesizing the diverse population of character strings, which diverse population comprises the library of biological polymers.

83. The method of claim 82, wherein the alteration comprises recombination of the pre-existing character strings.

5 **84.** The method of claim 82, wherein the biological polymers are selected from nucleic acids, polypeptides and peptide nucleic acids.

85. The method of claim 82, further comprising selecting members of the library of biological polymers for one or more activity.

10 **86.** The method of claim 85, further comprising filtering an additional library or an additional set of character strings by subtracting the additional library or the additional set of character strings with members of the library of biological polymers which display activity below a desired threshold.

15 **87.** The method of claim 85, further comprising filtering an additional library or an additional set of character strings by biasing the additional library or the additional set of character strings with members of the library of biological polymers which display activity above a desired threshold.

20 **88.** An integrated system comprising a computer having a first data set comprising a first character string, a second data set comprising a second character string, software for aligning the first and second character strings, software for performing a genetic operation on the first or second character string, an output file comprising a third data set comprising a third character string, the third character string comprising character string subsequences from the first and second character strings, and an oligonucleotide sequence output file comprising a plurality of overlapping oligonucleotide sequences corresponding to the third character string.

25 **89.** The integrated system of claim 88, the system further comprising an oligonucleotide synthesis machine for synthesizing the plurality of overlapping oligonucleotides.

90. The integrated system of claim 88, further comprising a plurality of oligonucleotides encoded by the plurality of overlapping oligonucleotide sequences, which oligonucleotides, when incubated in one or more cycles of chain extension, produce a third nucleic acid encoded by the third character string

5 **91.** The integrated system of claim 88, wherein the system further comprises a program with an instruction set for applying one or more genetic operator to the first or second character string, or to any other character string

10 **92.** The integrated system of claim 88, wherein the system further comprises a program with an instruction set for applying one or more genetic operator to the first or second character string, or to any other character string, wherein the genetic operator is selected from: a mutation, a multiplication, a fragmentation of the string or strings, a crossover between one or more strings, a ligation of strings, an elitism calculation, an alignment, a calculation of sequence homology or sequence similarity, a recursive use of one or more genetic operator for evolution of character strings, randomness, a deletion mutation, an insertion mutation, and
15 death.

93. A method of producing recombinant nucleic acids, the method comprising:
 providing two or more parental nucleic acid sequences;
 selecting cross-over sites for recombination between the two or more parental nucleic acid sequences, thereby defining one or more recombinant nucleic acids that result from a
20 cross-over between at least two of the two or more parental nucleic acids;
 determining a recombinant sequence for at least one of the one or more recombinant nucleic acids;
 selecting the at least one recombinant sequence in silico for one or more expected activity; and,
25 synthesizing the at least one recombinant sequence.

94. The method of claim 93, further comprising selecting bridging oligonucleotides which correspond to the cross-over sites.

95. The method of claim 94, wherein synthesizing the at least one recombinant sequence comprises providing fragments of the two or more parental nucleic acids and at least

one of corresponding bridge oligonucleotides, hybridizing the fragments and the bridge oligonucleotides and elongating the hybridized fragments with a polymerase or a ligase.

96. The method of claim 93, wherein the two or more parental sequences display low sequence similarity.

5 **97.** The method of claim 93, wherein selecting the at least one recombinant sequence in silico comprises one or more of:

- (i) performing an energy minimization analysis of the at least one recombinant sequence;
- (ii) performing a stability analysis of the at least one recombinant sequence;
- 10 (iii) comparing an energy minimized model of the at least one recombinant sequence to an energy minimized model of one or more of the two or more parental nucleic acids;
- (iv) performing protein threading on one or more encoded protein; and,
- (v) selecting the cross-over sites for recombination between the two or more
- 15 parental nucleic acid sequences to occur within regions of structural overlap, thereby determining the sequence of the at least one recombinant nucleic acid;
- (vi) performing one or more of: PDA, a branch-and-terminate a combinatorial optimization analysis, a dead end elimination, a genetic or mean-field analysis, or analysis of protein folding by threading, of the at least one recombinant
- 20 sequence;
- (vii) performing PDA of at least one of the two or more parental sequences; or
- (viii) comparing a PDA of the at least one recombinant sequence to a PDA of at least one of the two or more parental sequences.

98. The method of claim 93, wherein the step of selecting cross-over sites for

25 recombination between the two or more parental nucleic acid sequences and the step of selecting the at least one recombinant sequence in silico are performed simultaneously.



A DOCPHOENIX

APPL PARTS

_____**IMIS**_____
Internal Misc. Paper

_____**LET.**_____
Misc. Incoming Letter

_____**371P**_____
PCT Papers in a 371 Application

_____**A...**_____
Amendment Including Elections

_____**ABST**_____
Abstract

_____**ADS**_____
Application Data Sheet

_____**AF/D**_____
Affidavit or Exhibit Received

_____**APPENDIX**_____
Appendix

_____**ARTIFACT**_____
Artifact

_____**BIB**_____
Bib Data Sheet

_____**CLM**_____
Claim

_____**COMPUTER**_____
Computer Program Listing

_____**CRFL**_____
All CRF Papers for Backfile

_____**DIST**_____
Terminal Disclaimer Filed

_____**DRW**_____
Drawings

_____**FOR**_____
Foreign Reference

_____**FRPR**_____
Foreign Priority Papers

_____**IDS**_____
IDS Including 1449

_____**NPL**_____
Non-Patent Literature

_____**OATH**_____
Oath or Declaration

_____**PET.**_____
Petition

_____**RETMAIL**_____
Mail Returned by USPS

_____**SEQLIST**_____
Sequence Listing

_____**SPEC**_____
Specification

_____**SPEC NO**_____
Specification Not in English

_____**TRNA**_____
Transmittal New Application

_____**CTNF**_____
Count Non-Final

_____**CTRS**_____
Count Restriction

_____**EXIN**_____
Examiner Interview

_____**M903**_____
DO/EO Acceptance

_____**M905**_____
DO/EO Missing Requirement

_____**NFDR**_____
Formal Drawing Required

_____**NOA**_____
Notice of Allowance

_____**PETDEC**_____
Petition Decision

OUTGOING

_____**CTMS**_____
Misc. Office Action

_____**1449**_____
Signed 1449

_____**892**_____
892

_____**ABN**_____
Abandonment

_____**APDEC**_____
Board of Appeals Decision

_____**APEA**_____
Examiner Answer

_____**CTAV**_____
Count Advisory Action

_____**CTEQ**_____
Count Ex parte Quayle

_____**CTFR**_____
Count Final Rejection

INCOMING

_____**AP.B**_____
Appeal Brief

_____**C.AD**_____
Change of Address

_____**N/AP**_____
Notice of Appeal

_____**PA..**_____
Change in Power of Attorney

_____**REM**_____
Applicant Remarks in Amendment

_____**XT/**_____
Extension of Time filed separate

Internal

_____**SRNT**_____
Examiner Search Notes

_____**CLMPTO**_____
PTO Prepared Complete Claim Set

_____**ECBOX**_____
Evidence Copy Box Identification

_____**WCLM**_____
Claim Worksheet

_____**WFEE**_____
Fee Worksheet

File Wrapper

_____**FWCLM**_____
File Wrapper Claim

_____**IIFW**_____
File Wrapper Issue Information

_____**SRFW**_____
File Wrapper Search Info

BACKFILE DOCUMENT INDEX SHEET

**METHODS FOR MAKING CHARACTER STRINGS, POLYNUCLEOTIDES AND
POLYPEPTIDES HAVING DESIRED CHARACTERISTICS**

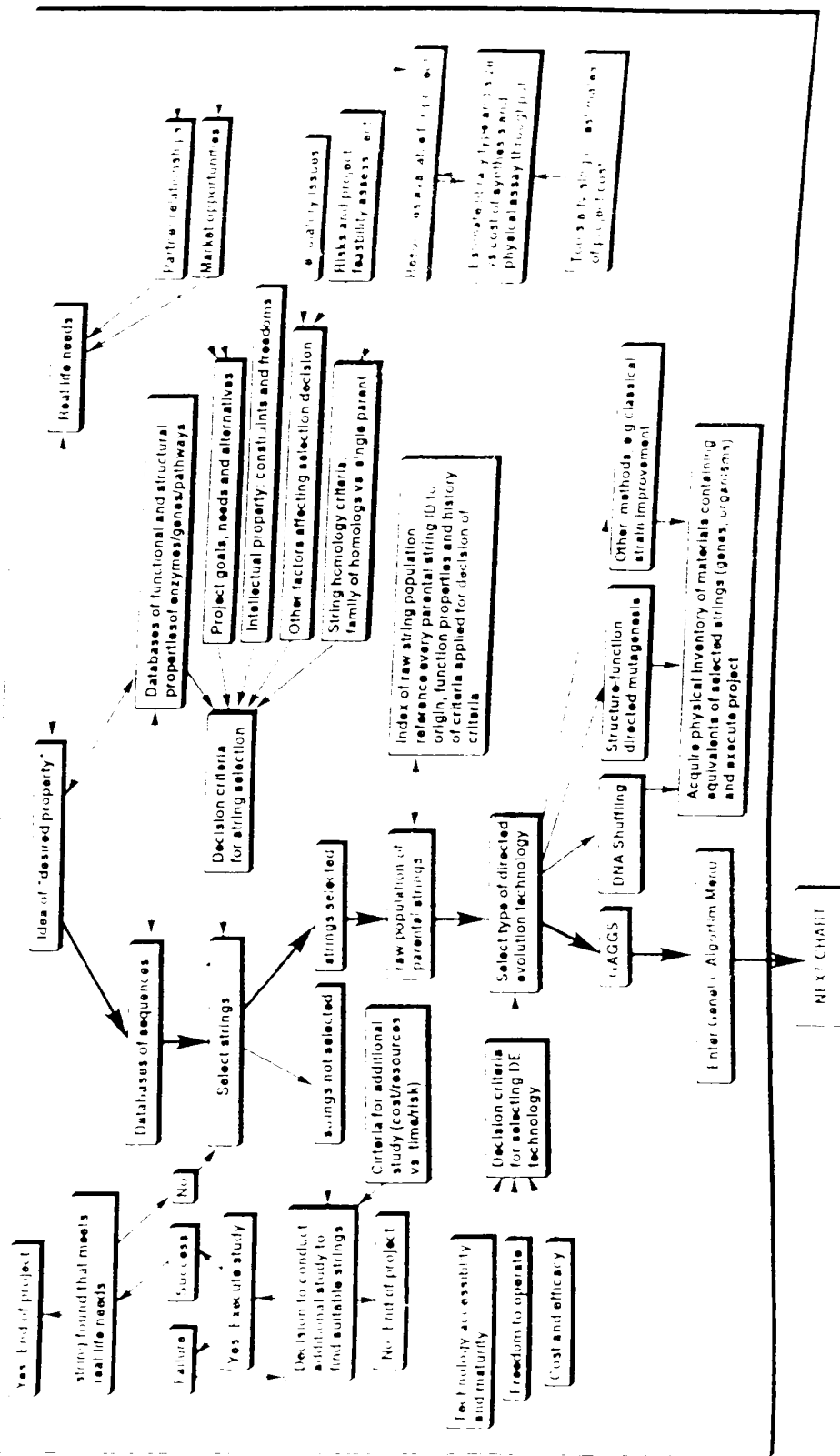
ABSTRACT OF THE DISCLOSURE

5

"In silico" nucleic acid recombination methods, related integrated systems
utilizing genetic operators and libraries made by in silico shuffling methods are provided.

ma289-4.app.doc

Figure 1





A DOCPHOENIX

APPL PARTS

_____**IMIS**_____
Internal Misc. Paper

_____**LET.**_____
Misc. Incoming Letter

_____**371P**_____
PCT Papers in a 371 Application

_____**A...**_____
Amendment Including Elections

_____**ABST**_____
Abstract

_____**ADS**_____
Application Data Sheet

_____**AF/D**_____
Affidavit or Exhibit Received

_____**APPENDIX**_____
Appendix

_____**ARTIFACT**_____
Artifact

_____**BIB**_____
Bib Data Sheet

_____**CLM**_____
Claim

_____**COMPUTER**_____
Computer Program Listing

_____**CRFL**_____
All CRF Papers for Backfile

_____**DIST**_____
Terminal Disclaimer Filed

_____**DRW** 14_____
Drawings

_____**FOR**_____
Foreign Reference

_____**FRPR**_____
Foreign Priority Papers

_____**IDS**_____
IDS Including 1449

_____**NPL**_____
Non-Patent Literature

_____**OATH**_____
Oath or Declaration

_____**PET.**_____
Petition

_____**RETMAIL**_____
Mail Returned by USPS

_____**SEQLIST**_____
Sequence Listing

_____**SPEC**_____
Specification

_____**SPEC NO**_____
Specification Not in English

_____**TRNA**_____
Transmittal New Application

_____**CTNF**_____
Count Non-Final

_____**CTRS**_____
Count Restriction

_____**EXIN**_____
Examiner Interview

_____**M903**_____
DO/EO Acceptance

_____**M905**_____
DO/EO Missing Requirement

_____**NFDR**_____
Formal Drawing Required

_____**NOA**_____
Notice of Allowance

_____**PETDEC**_____
Petition Decision

OUTGOING

_____**CTMS**_____
Misc. Office Action

_____**1449**_____
Signed 1449

_____**892**_____
892

_____**ABN**_____
Abandonment

_____**APDEC**_____
Board of Appeals Decision

_____**APEA**_____
Examiner Answer

_____**CTAV**_____
Count Advisory Action

_____**CTEQ**_____
Count Ex parte Quayle

_____**CTFR**_____
Count Final Rejection

INCOMING

_____**AP.B**_____
Appeal Brief

_____**C.AD**_____
Change of Address

_____**N/AP**_____
Notice of Appeal

_____**PA..**_____
Change in Power of Attorney

_____**REM**_____
Applicant Remarks in Amendment

_____**XT/**_____
Extension of Time filed separate

Internal

_____**SRNT**_____
Examiner Search Notes

_____**CLMPTO**_____
PTO Prepared Complete Claim Set

_____**ECBOX**_____
Evidence Copy Box Identification

_____**WCLM**_____
Claim Worksheet

_____**WFEE**_____
Fee Worksheet

File Wrapper

_____**FWCLM**_____
File Wrapper Claim

_____**IIFW**_____
File Wrapper Issue Information

_____**SRFW**_____
File Wrapper Search Info

BACKFILE DOCUMENT INDEX SHEET

100

101

102

103

104

105

106

107

108

109

110

111

112

113

114

115

116

117

118

119

120

121

122

123

124

125

126

127

128

129

130

131

132

133

134

135

136

137

138

139

140

141

142

143

144

145

146

147

148

149

150

151

152

153

154

155

156

157

158

159

160

161

162

163

164

165

166

167

168

169

170

171

172

173

174

175

176

177

178

179

180

181

182

183

184

185

186

187

188

189

190

191

192

193

194

195

196

197

198

199

200

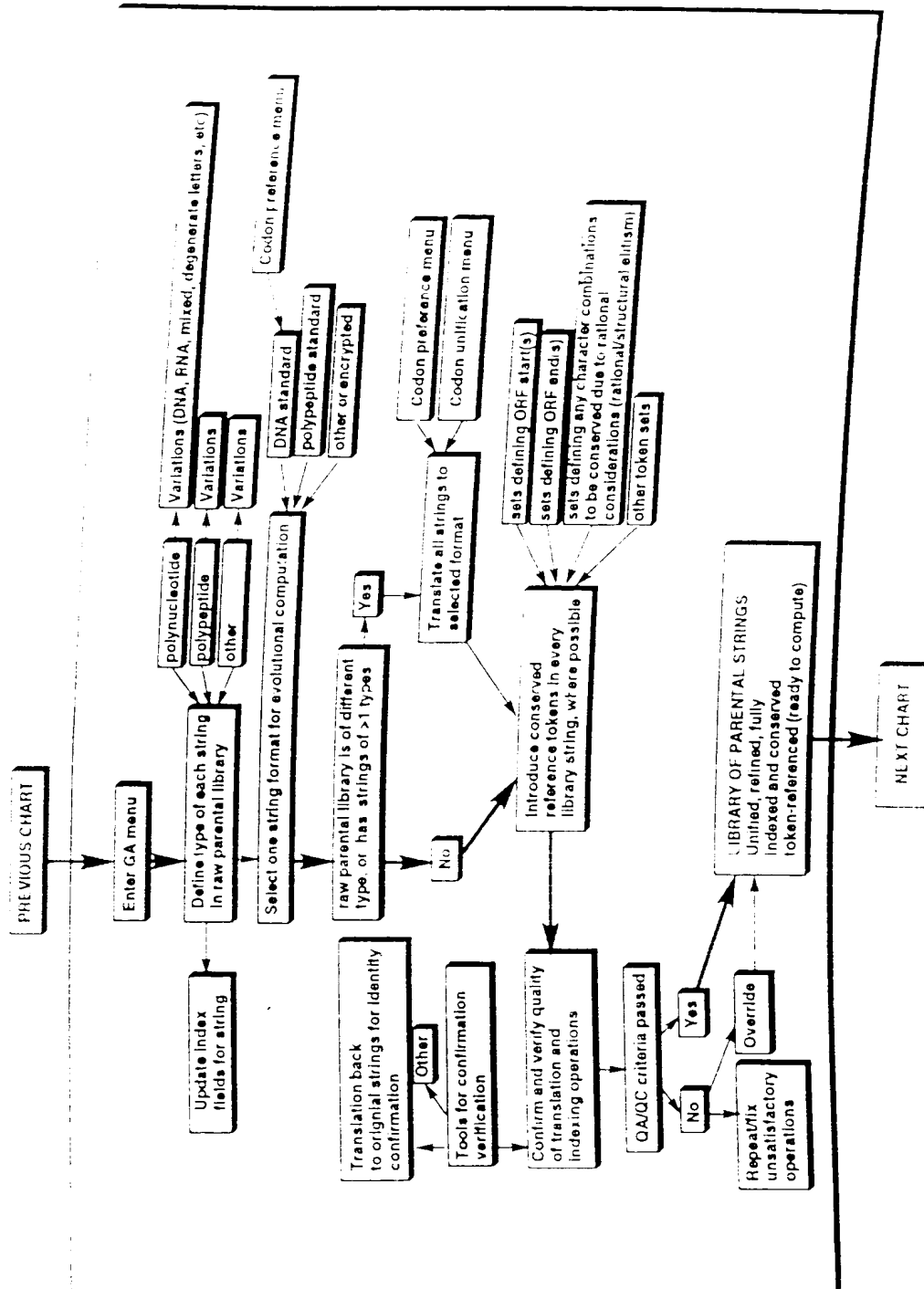


Figure 2

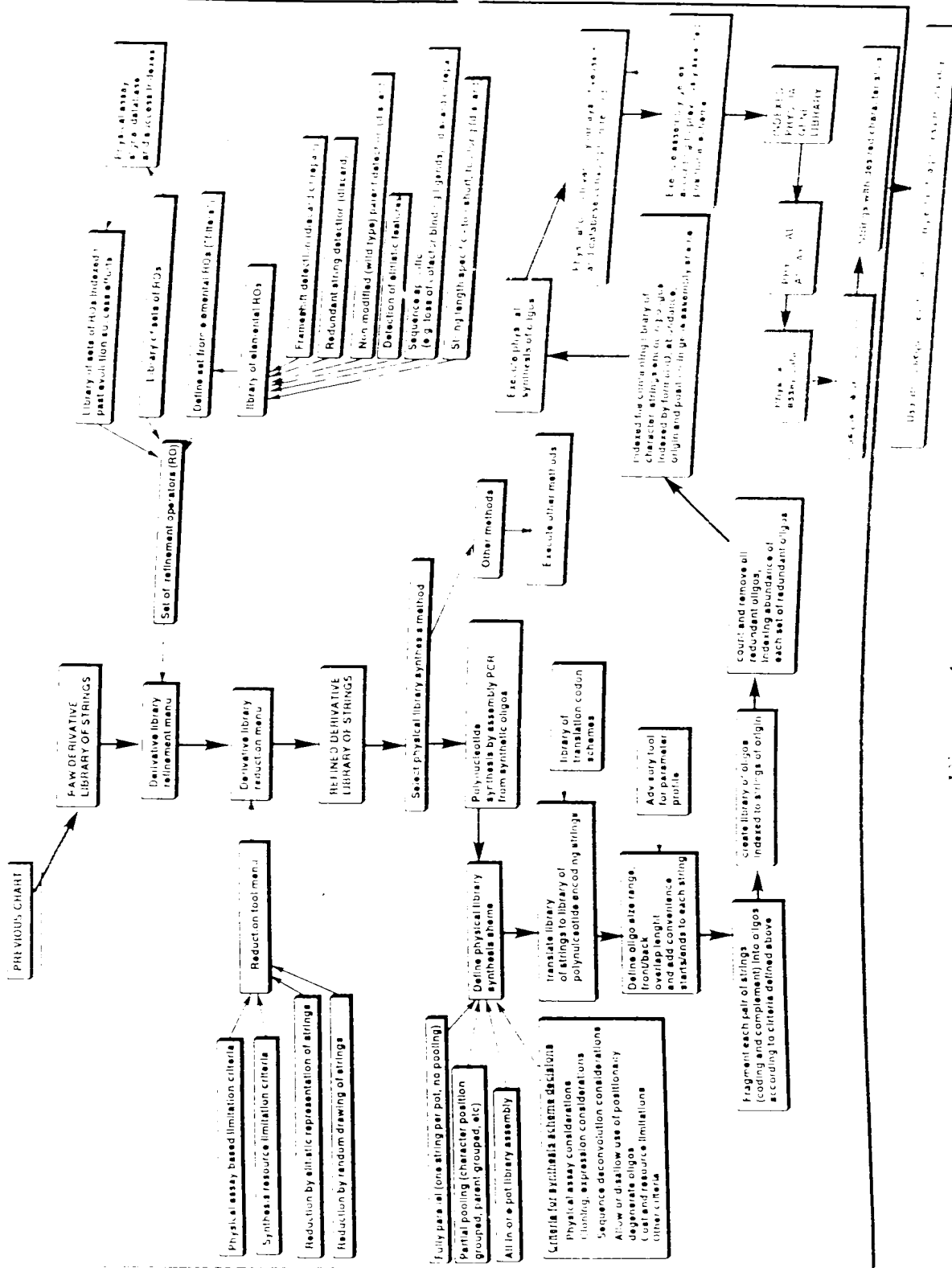


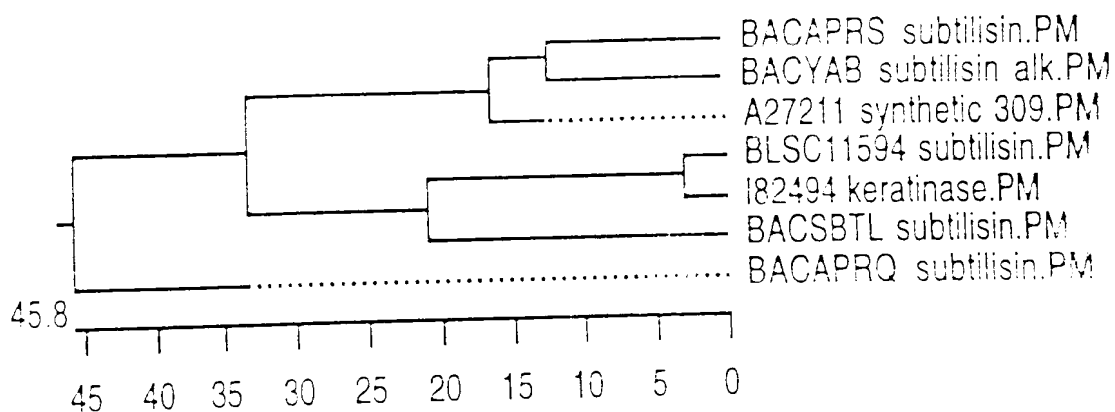
Figure 4

Figure 5

Percent Similarity

	1	2	3	4	5	6	7	
1		62.1	81.4	57.6	81.8	56.1	59.1	1
2	50.5		61.0	54.9	59.5	58.2	60.8	2
3	21.0	52.0		54.6	78.4	50.6	53.2	3
4	54.4	63.3	62.3		52.0	64.6	67.9	4
5	20.5	54.9	25.1	65.6		53.9	56.5	5
6	58.6	56.6	72.2	44.2	63.4		94.9	6
7	52.5	51.4	66.0	38.5	57.8	4.9		7
	1	2	3	4	5	6	7	

A27211 synthetic 309.PM
 BACAPRQ subtilisin.PM
 BACAPRS subtilisin.PM
 BACSBTL subtilisin.PM
 BACYAB subtilisin alk.PM
 BLSC11594 subtilisin.PM
 I82494 keratinase.PM



UNACCEPTED

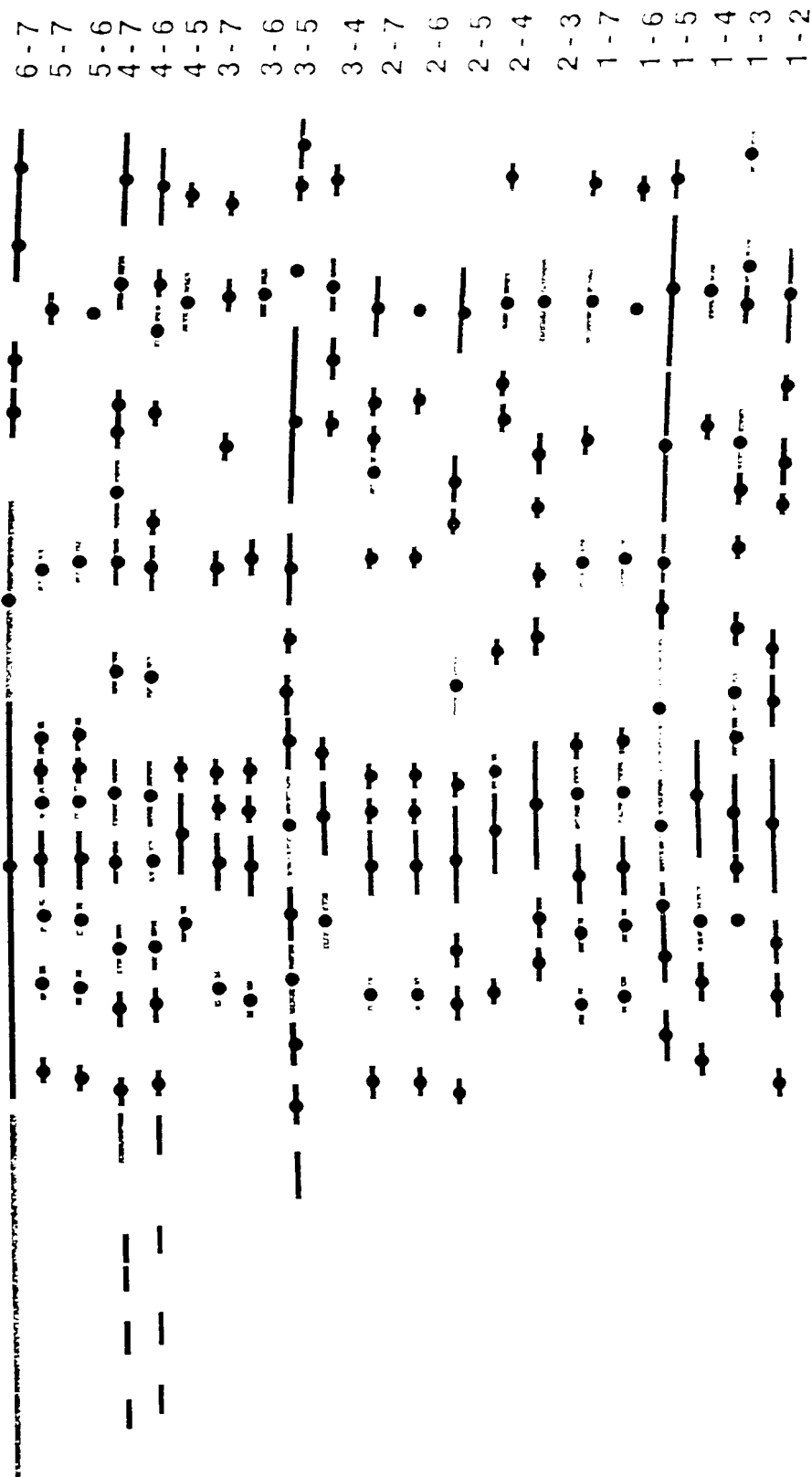


Figure 7

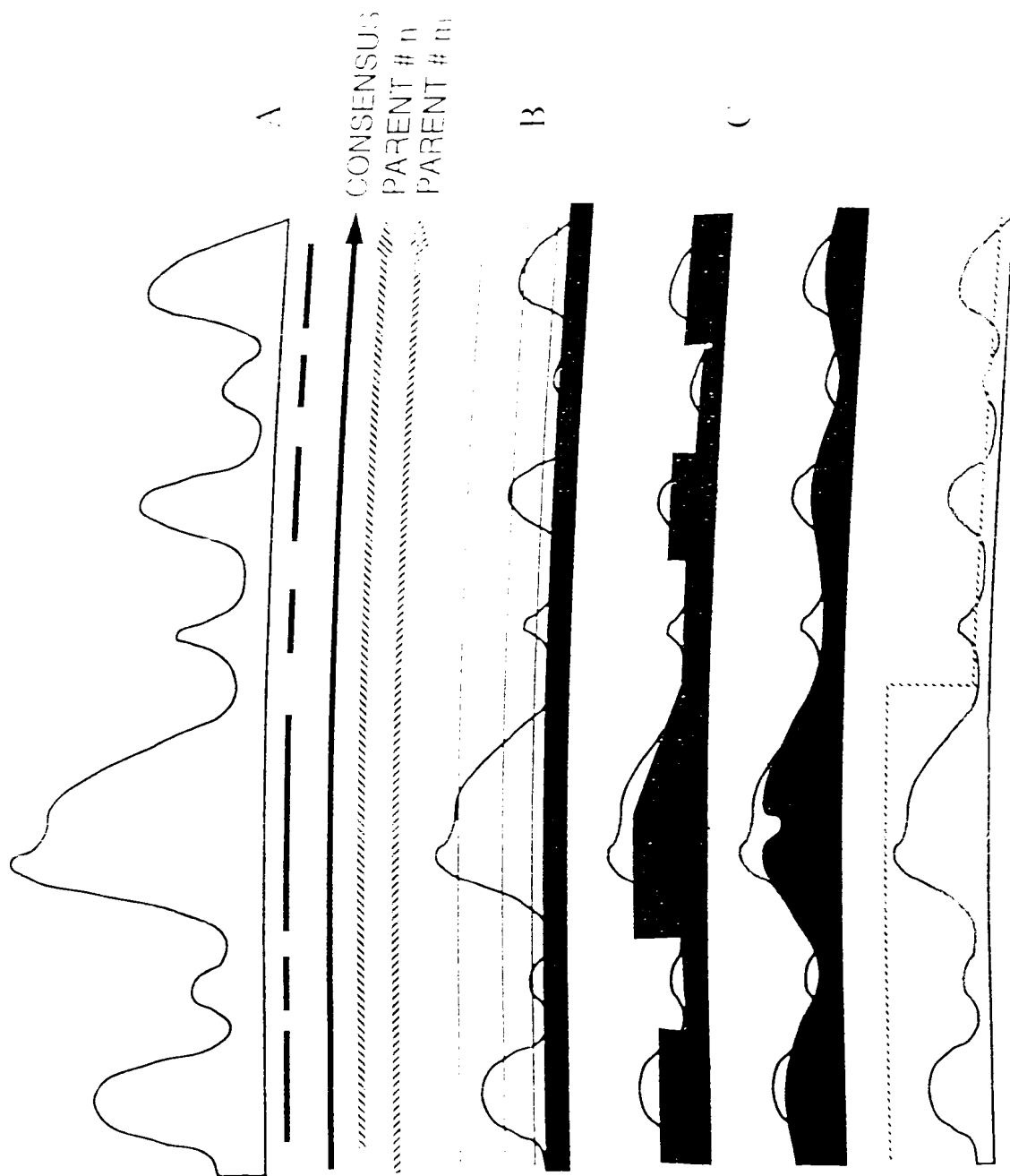


Figure 8

Figure 9

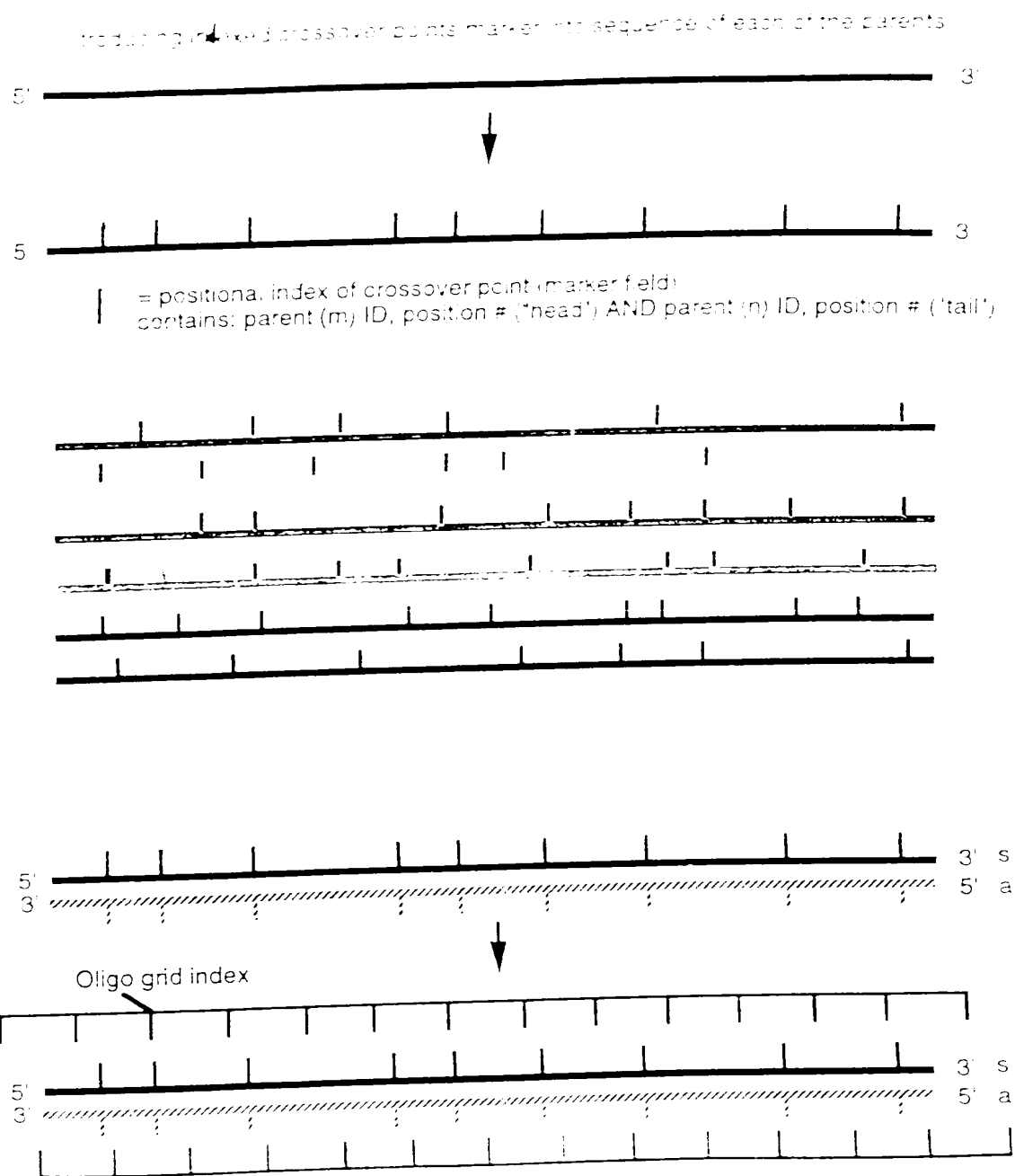
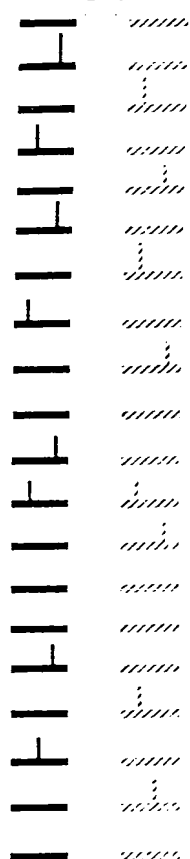


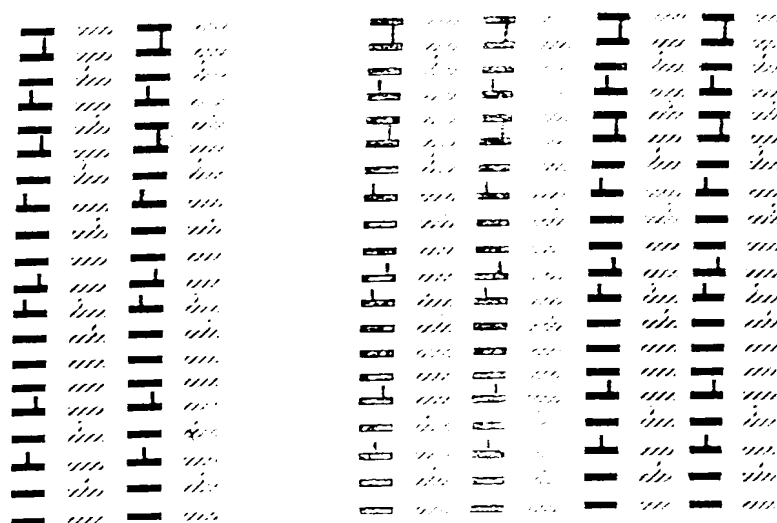
Figure 10

SET OF OLIGOS
TO ASSEMBLE A PARENT

5 3 5 3



COMPLETE INVENTORY OF
OLIGO SEQUENCES
(TO ASSEMBLE ALL PARENTS)



FIND ALL PAIRS OF PAIRS OF
OLIGO SEQUENCES WITH
MATCHING PAIRWISE
CROSSOVER INDEXES

SUB-INVENTORY OF OLIGOS WITH
CROSSOVER MARKERS

11 15

Figure 11

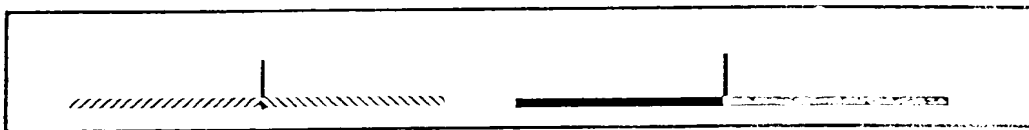
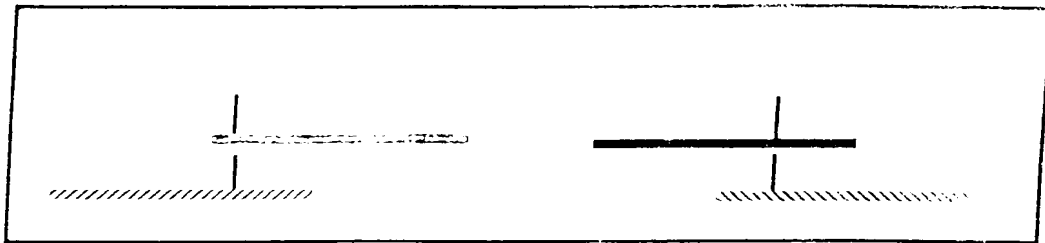
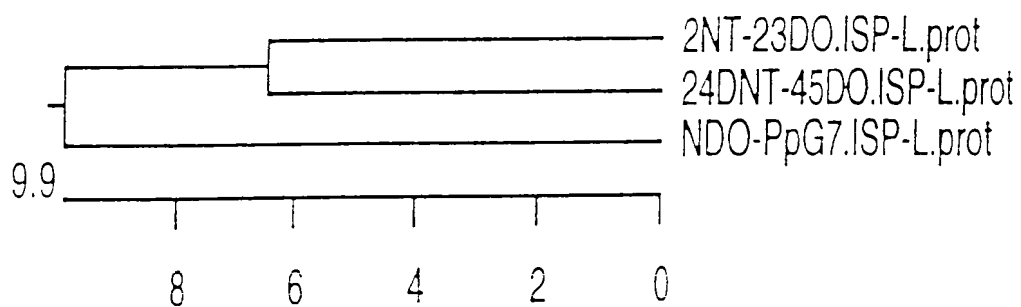


Figure 12

Percent Similarity

	1	2	3	
1		87.1	79.1	1
2	13.0		83.3	2
3	21.7	17.9		3
	1	2	3	

24DNT-45DO.ISP-L.prot
2NT-23DO.ISP-L.prot
NDO-PpG7.ISP-L.prot



13 15

Figure 13

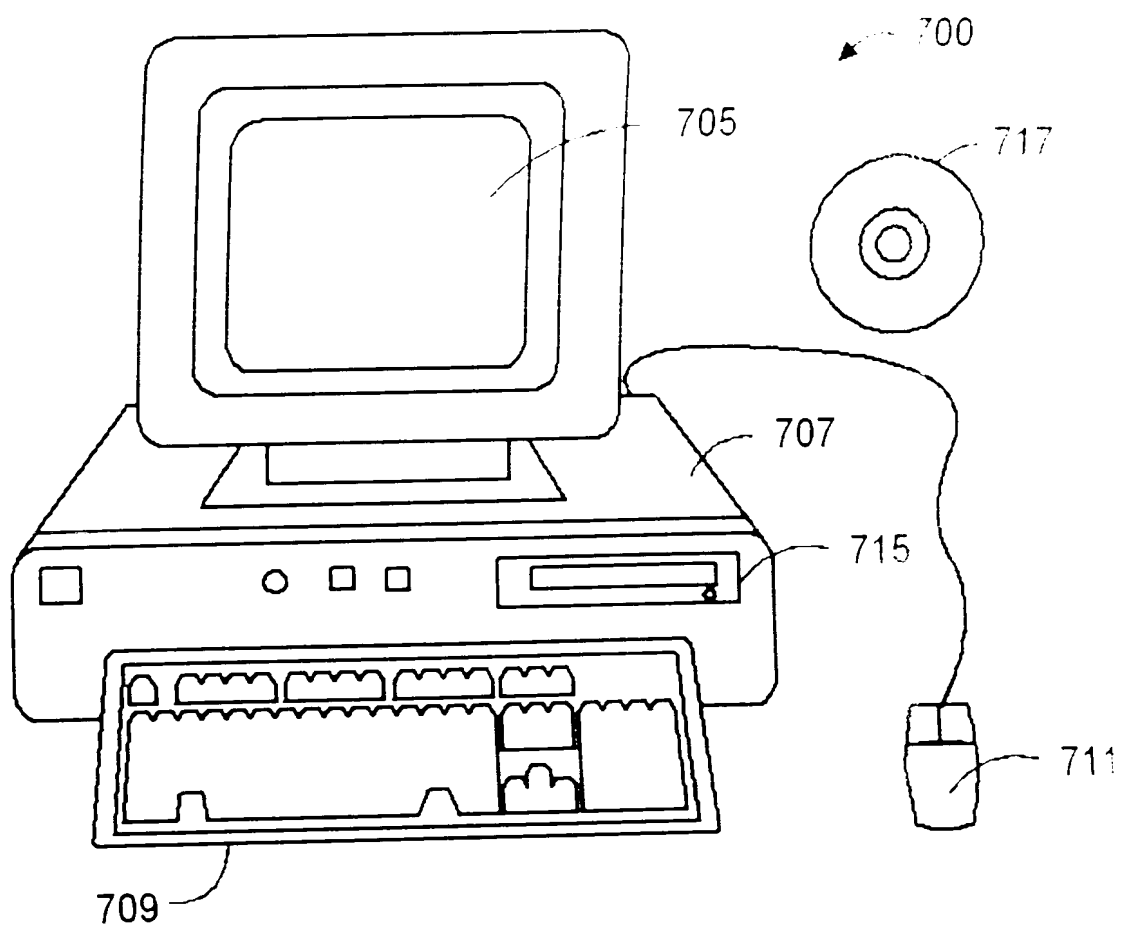
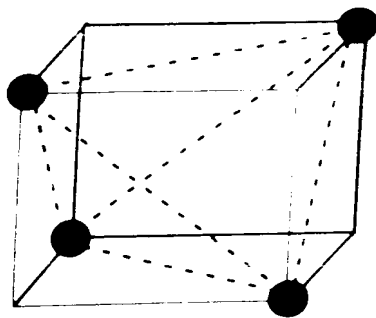


Figure 14



Geometric representation
Equal distance between
each nucleotide

A	-1	-1	+1
C	+1	-1	-1
G	-1	+1	-1
T	+1	+1	+1

Numeric representation

15 15

Figure 15

